數學傳播 49卷1期, pp. 43-49

機率學者看量子力學 (二)

謝南瑞

前言: 本文是參考資料 [4] 的後續; 討論量子量測 (quantum measurements) 的機率觀點。文章的格調, 和前者一致, 除了前言和結語外, 分成一些小節, 講述相關議題。和前者不同的是, 爲了編排打字的方便, 作者不再使用下橫線或其他字體來註明重要的名詞; 值得讀者關注的名詞, 一律以中文附英文的方式呈現, 希望讀者自行依此, 在中文或英文維基百科上, 或是 Google 搜尋引擎找尋。又, 人名的呈現方式, 和前文 [4] 一樣, 除 "衆所皆知" 的用中文外, 餘以英文表達, 且附生卒年份, 和最著名的學術成就。

波爾(Neals Bohr, 1885∼1962, 量子力學先驅, 1922年諾貝爾物理學獎得主) 曾有句名言: 若無量測, 無事物存在 (Nothing exists, until it is measured)。當然, 這句話有其適用範圍, 愛因斯坦就曾以月亮的存在, 反駁這句話。作者往昔, 在講述一些課程時, 喜歡用 MPST 表達人類的四大文明支柱, 此爲數學哲學科學神學四個英文字的起頭字母。當然, 人類也有人文藝術和科技, 但是作者認爲, 後三者都是前四者的衍生品 (derivatives)。在 MPST 中, 作者認爲: 科學絕對須以量測爲基石, 哲學神學則絕對不能, 數學是介於可和不可與須和不須之間。

以下, 僅就一個機率學者的角度, 對量子力學的量測提出討論。必須說明: 量子量測, 有一個古典量測沒有的特殊性。即是它的量測問題 (measurements problems in quantum mechanics)。

(一) 由古典量測到量子量測

在近代科學的進展中, 最著名的量測, 應該是加利略在比薩斜塔上, 丟下一大一小的兩個球, 量測兩者是否同時著地。古典量測, 可以分成確定性量測和不確定性量測兩類。確定性的量測, 是我們最常遇見的, 小至把一個桌上的蘋果, 拿來稱它的重量, 大至天文學者測量彗星的軌跡。這些量測, 有個共通點: 量測前和量測後, 我們"看不出"受測物體有什麼改變! 至於不確定的量測, 我們簡述如下:

某一隨機量, 我們希望透過量測, 來了解此隨機量的機率分佈 (probability distribution)。

有兩種主要方式: (1) 假設, 此隨機量可以獨立且同分佈地 (independently and identically distributed) 實現 N 次, 而 N 夠大。則大數法則 (law of large numbers) 告訴我們, 此 N

次實現的樣本平均值, 趨近於某一常值, 即此隨機量的“期望值”。而, 中央極限定理 (central limit theorem) 可以用來估計上述趨近的誤差。(2) 假設, 此隨機量不能以 (1) 方式作夠多次的實現, 我們引入條件機率 (conditional probabilities) 的觀念。我們設此隨機量 X, 且已知某一事件 B 已發生, 在此情況下, 重新估計 X 的機率分佈, 通常以 $E[X \mid B]$ 表達, 稱爲 X 對 B 的條件分佈 (conditional distribution), 此觀念在一般的機率統計教科書中, 都有詳述; 最重要的實用式, 是貝氏公式 (Bayes formula); 讀者可以參考相關教科書。在數學傳播上, 也早已有若干文章討論條件機率。

以上的不確定性量測, 仍然是有確定性量測屬性: 量測前和量測後, 我們只是作受測的隨機性之機率分佈的重估; 此受測隨機性的機率分佈重估, 例如所丟的骰子是否代表一個公正的骰子; 我們懷疑某賭場所用骰子, 出現小點數的機率大於出現大點數的機率, 而使用假說檢定 (testing hypothesis), 這目的和所用的架構, 量測前和量測後, 沒有改變。

然而, 量子量測是截然不同! 我們以單獨的小節討論它。

(二) 量子量測的特殊性和解釋

首先, 我們以薛丁格的貓 (Schrödinger’s cat) 爲例。利用 [4] 的符號, 這是佈於複數系 $\mathbb{C}$ 上的 $n=2$維向量空間, 而每個量子態爲 2 維單位長之向量, 而貓是疊加態

$$\frac{1}{\sqrt{2}}\vec{v}_1 + \frac{1}{\sqrt{2}}\vec{v}_2, \quad \vec{v}_1 = (1,0), \quad \vec{v}_2 = (0,1).$$

上式右方第一項爲生態而第二項爲死態, 是以, 上式代表貓處於由兩態以等係數 (意謂, 等機率) 構成的量子疊加態。若觀測者不打開盒子, 則此量子態持續。但是, 若觀測者打開盒子, 則看到貓要嘛是生要嘛是死, 量子態消失了! 量子系一碰上觀測或量測, 則量子態立即消失而成爲一個巨觀態且無法回復, 此現象和前節的古典量測完全不同! 物理學者稱爲量子坍塌(quantum collapse)。量子量測的坍塌現象是量子力學和古典力學的主要不同之一。

如何解釋? 以下是三個被普遍接受的看法:

(A) 這不用解釋, 物理是講究實測數據, 所以就是: 閉嘴, 好好算! (Shut up and calculate!) (註: 這話經常被認爲是理查費曼所說, 但是實際上出自 David Mermin (1935∼) 所言)。

(B) 隱變量理論 (Hidden-variable theory), 最早似應是出自愛因斯坦。他認爲量子力學若是完備的 (complete), 則必須系統內存在某個隱含的變量, 而使量測結果是對此變量的“平均”值, 參考其著名論文 [2]。此變量必須是局域 (locality), 且可以解釋 [2] 中所提出的思想實驗(thought experiment) 出現之超局域現象 (non-locality)。然而此變量的本質, 似乎迄今無法理解。

(C) 多世界解釋 (Many-worlds interpretation), 最早出自 Hugh Everett III (1930∼1982)。

於 1970 年代後逐漸被重視。著名的美國科普作家 Sean M. Carroll (1966∼) 有本書 [1] 以此爲主題。此理論利用量子糾纏 (quantum entanglement): 若是觀測者觀測到貓是活的, 則觀測者和活貓形成一個糾纏態, 否則觀測者和死貓形成一個糾纏態; 亦即, 在觀測者打開貓盒的剎那, 時空產生分岔 (bifurcation) 而爲兩個 "世界", 這兩個世界平行地存在, 但是無法交通。

(三) 量子量測的 Born 法則

我們在此節, 將前文 [4] 中的 Born 法則, 作一推廣, 如下。

先作回顧: 我們設量子系爲含有 n 個量子態的系統。因此, 其數學架構爲佈於複數系 $\mathbb{C}$ 上的 n 維向量空間, 而每個量子態爲單位長之向量 $\vec{v} = (v_1, \ldots, v_n), \|\vec{v}\|^2 = |v_1|^2 + \cdots + |v_n|^2 = 1$。由量子態 $\vec{v}$ 躍遷到量子態 $\vec{w}$ 的機率振幅 (probability amplitude) 是內積 $(\vec{v}, \vec{w}) = \sum_{i=1}^n v_i \overline{w_i}$ 的平方。一個可觀測量 (observable), 是一個 n 階自伴 (self-adjoint) 方陣: $A = [a_{ij}]_{i,j=1}^n$, $a_{ij} \in \mathbb{C}$, $A = \overline{A}^t$, A^t 表 A 的轉置方陣, 正橫槓表示取共軛; 亦即,

$$[a_{i,j}] = [\overline{a_{j,i}}].$$

可證明:

若 $\lambda \in \mathbb{C}$ 爲 A 的一個非零特徵值, 即存在某非零 $\vec{v} \in \mathbb{C}^n$ 使 $A\vec{v} = \lambda\vec{v}$ ($\vec{v}$ 作一轉置), 則 λ 必爲實數。現在, 我們證明

Born **法則**: 設可觀測量 A 有 k 個相異的非零特徵值 λ_j, $j = 1, \ldots, k$, $k < n$, 每個 λ_j 相應的特徵向量線性子空間 (linear subspace spanned by the orthonormal eigenvectors corresponding to λ_j) 爲 V_j。則, 可觀測量 A 可分解成

$$A = \sum_{i=1}^k AP_i, \qquad P_i : \mathbb{C}^n \to V_i,$$

對每個量子態 $\vec{v}$, 以 A 量測 $\vec{v}$, 必得到某 λ_i, 且量測到 λ_i 的機率是

$$|(\vec{v}, \vec{v}_i)|^2;$$

上式中, P_i 表 $\mathbb{C}^n$ 至 V_i 的射影映射 (projection mapping):

$$P_i(\vec{v}) = \vec{v}_i := \min\{|\vec{v} - \vec{v'}| : \vec{v'} \in V_i\}.$$

證明: 設線性子空間 V_i 的維度爲 $\dim(V_i)$, 令 $\dim(V_i) = n_i$。則有限維度譜表現定理 (finite-dimensional spectral representation theorem) 告訴我們: $\sum_{i=1}^k n_i = n$。因此, 若以矩陣

表示射影映射, 我們有以下的矩陣和:

$$\sum_{i=1}^{k} P_i = I_n,$$

其中 I_n 表示 $n \times n$ 的單位矩陣。上述公式表明了矩陣 A 的分解。至於觀測到特徵值 λ_i 的機率, 我們注意到:

$$\big(A(\vec{v}),\vec{v}\big) = \Big(\big(\sum_{i=1}^{k} AP_i\big)\vec{v},\vec{v}\Big) = \sum_{i=1}^{k}\Big(A\big(P_i\vec{v}\big),\vec{v}\Big) = \sum_{i=1}^{k}(\vec{v}_i, A(\vec{v})) = \sum_{i=1}^{k}\lambda_i(\vec{v},\vec{v}_i).$$

其中 $\vec{v}_i = P_i\vec{v}$。 Q.E.D.

注意: 若是 $k = n$, 則每個 λ_j 相應的特徵向量子空間維度爲 1, 此時, 以上定理即是前文 [4] 中所予的。

(四) 量子量測的一個機率架構

在此節, 我們討論一個作者認爲有興趣研讀量子力學的數學理論之讀者, 值得進行的架構。此節取材自 [3] 第一章。

由前文 [4] 和上述的 Born 法則, 每個 n 質點量子系中的量子態, 經一個可觀測量, 定義成 n 階自伴方陣 A, 測量所得, 是 A 的某個特徵值, 且附隨有觀測到此值的機率。那麼, 我們何不把量測, 看成將某個 "態之集合" 中的每一個 "態", 映射到某個"值之集合"上的機率分佈? 以下是此構想的形式定義 : 我們以草寫 $\mathcal{U}$ 代表"態之集合", 而每個態以正寫 S 表達; 又, 以正寫 U 表"值之集合", 而 $P(U)$ 表 U 上機率分佈之全體 (註: 爲了方便理解, 我們可以設 U 爲有限集或可列集, 是以 $P(U)$ 的成員, 可表成 p_i, $i = 1,\ldots,k$ 或 $i = 1,2,\ldots$, 而每個 p_i 大於或等於 0, 且對 i 相加總和爲 1)。我們另須要兩個觀念。一個是凸集合 (convex set): 一集合 A 是凸的, 若對任何 A 中兩成員 t,u 和任何數 a, $0 < a < 1$, 凸組合 (convex combination) $at + (1-a)u$ 也在 A 中。顯然, $P(U)$ 是凸集, 我們須設態之集合 $\mathcal{U}$ 也是凸的。另一個觀念是仿射映射 (affine mapping): 由集合 A 至集合 B 的一個映射 F 爲仿射, 若是 $F(u) = L(u) + b$; 而 L 是一個線性變換, 且 b 是一個常值; 當然, 我們須設集合 A, B 有適當的加性和線性結構。以下是我們對量測一詞的定義:

定義: 一個態之集合 $\mathcal{U}$ 的量測 (measurement), 是一個由 $\mathcal{U}$ 至 $P(U)$, U 爲量測所能得到數值的全體, 的仿射映射 F。

由仿射性, 可以看出, 任予 k 個態 $S_1,\ldots,S_k$, 和一組權重 (weights) $w_1,\ldots,w_k$ (即, w_i, $i = 1,\ldots,k$, w_i 大於或等於 0, 且對 i 相加和爲 1), F 將由 $\{w_i\}$ 和 $\{S_i\}$ 所形成的混

合態 (mixture state) 映射至由 $\{w_i\}$ 和 $\{F(S_i)\}$ 所成的線性組合, 即

$$F\Big(\sum_{i=1}^{k} w_i S_i\Big) \to \sum_{i=1}^{k} w_i F(S_i).$$

對於一個凸集 A, A 的極點 (extreme point) 集, 是一個子集 A_1, 而使 A 中每個成員, 皆可表成 A_1 中成員的凸組合, 即

$$\forall\, t \in A, \quad \exists u_1, \ldots, u_k \in A_1, \quad s.t.$$

$$t = \sum_{i=1}^{k} a_i u_i, 1 < a_1, \ldots, a_k < 1, \quad \sum_{i=1}^{k} a_i = 1.$$

刻畫一個凸集的所有極點, 是凸集合理論中的基本工作。

在討論本節的主要結果前, 我們先說明: 上述的量測定義, 也適用於古典量測。古典確定性量測, 態的全體爲某個屬性 (例如, 稱量蘋果) 之所有古典物理態, 而仿射 F 爲

$$F : S \;\to\; F(S) = I_{[m-e,m+e]}(S),$$

上式中, m 爲給定之測量值, 而 e 爲一誤差值。上式右邊, 代表一般稱爲 delta 型式的機率分佈, I 值爲 $=1$, 若態 S 之測量值在區間 $[m-e, m+e]$ 內; 值爲 $=0$, 若 S 之測量值不在此區間內 (若是 $e=0$, 則代表此量測全無誤差, 但此在實測上無可能)。古典不確定性量測, 態的全體爲所有古典隨機態 X, 而仿射 F 爲

$$F : X \;\to\; F(X) = P_X,$$

上式右邊, 代表 X 的機率分佈 (U 爲隨機態的所有可能值之集合)。

以下, 我們論述此節的主要結果。我們仍是只考慮 n 個質點所成的量子系, 所以, 其數學架構爲 $\mathbb{C}^n$, 佈於複數系上的 n 維向量空間, 而量子態以 $\mathbb{C}^n$ 中的單位長向量 $\vec{v}$ 表之。一個複數元 n 階方陣 A 稱爲正定的 (positive definite), 若是

$$\vec{v}A\Big(\overline{\vec{v}}\Big)^t \geq 0, \quad \forall\, \vec{v}.$$

記爲 $A \geq 0$。而 A 的對角線各元之和, 以 $\mathrm{Tr}\,(A)$ 表之, 即

$$\mathrm{Tr}\,(A) = a_{1,1} + \cdots + a_{n,n}.$$

依據量子力學之用語, 稱一個 n 階正定自伴方陣 A 爲一個密度矩陣 (density matrix), 若是 $\mathrm{Tr}\,(A) = 1$。可以證明: 每一量子態 $\vec{v}$ 所成的 n 階正定自伴方陣 $\langle \vec{v}, \vec{v}\rangle$ (回想: 符號 $\langle \vec{v}, \vec{v}\rangle$ 代表 $\vec{v}$ 和 $\vec{v}$ 自己的 "外積", 亦即, $\vec{v}$ 和其共軛轉置所成的 n 階自伴方陣) 必爲一密度矩陣。我們以草寫 $\mathscr{D}_n$ 表 n 階密度矩陣的全體。我們有以下的刻畫:

性質: $\mathscr{D}_n$ 是一凸集, 其極點的全體爲 $\langle\vec{v},\vec{v}\rangle$, $\vec{v}$ 爲 n 維單位長向量的全體。

我們不證明此性質, 讀者可以在線性代數的教本或是網站中找到其證明; [3] 第一章 (該章之性質 1.2.3) 內亦有。依據量子力學之用語, 每個 $\langle\vec{v},\vec{v}\rangle$ 皆是純態 (pure state)。

以下, 是本節的主要結果:

定理：一個 n 質點量子系的量子量測, 視爲態之集合 $\mathscr{D}_n$ 至 $P(U)$ 的仿射變換 F, 其中值之集合 U 爲有限集 $\{1,\ldots,k\}$, 可以唯一的被表達成

$$F(S)_i = \mathrm{Tr}\,(SM_i), \quad i\in U.$$

於上式, $\{M_i\}, i=1,\ldots,k$, 爲一組 n 階正定自伴矩陣, 且對 i 作和的加總爲 $I_{n,n}$ 階單位方陣; 以數學分析用語, $\{M_i\}$ 構成一個么元分解 (resolution of identity)。

證明：我們將證明分成四步驟:

1. 我們定義 $\mathscr{L}_n$ 爲所有 n 階自伴矩陣所構成的線性子空間 (這是一個在 $\mathbb{C}$ 上的有限維線性空間), 其由 $\mathscr{D}_n$ 中元素的所有實係數 (注意, 此爲重要) 線性組合所構成。接著, 我們將映射 F 的定義域 $\mathscr{D}_n$ 擴展到 $\mathscr{L}_n$, 具體定義如下: 對於

   $$T=\sum_{j=1}^{l}a_jA_j, \quad 令\ F(T)=\sum_{j=1}^{l}a_jT(A_j), \quad T(A_j)=F(A_j), \quad a_j\in\mathbb{R},\ A_j\in\mathscr{D}_n.$$

   首先, 我們證明這樣的擴展必然存在。根據有限維的譜定理 (應用於自伴矩陣), 我們可以將任意 n 階自伴矩陣表示爲一個具有實係數的分解 (回憶: 自伴矩陣的所有特徵值必爲實數), 我們也注意到任何投影算子都是正定的。我們將略過上述 F 的擴展定義在數學上是良好定義 (well-defined) 的證明。也就是說, 若 $\mathscr{L}_n$ 中的同一元素有兩種不同的表示方式, 則上述定義的 F 對這兩種表示必須給出相同的值。

2. 現在, 因爲 F 是定義在有限維線性空間 $\mathscr{L}_n$ 上的實數值 (注意, 此爲重要) 線性泛函, 且每個 $T\in\mathscr{L}_n$ 本身是 $n\times n$ 的矩陣, 記作 $T=[t_{i,j}]$, 根據 Riesz 表現定理 (Riesz Representation Theorem), 必存在一個 $n\times n$ 的矩陣 $M=[m_{k,l}]$ (具有複數元), 使得

   $$F(T)=\big[[M,T]\big]:=\sum_{j,k}t_{j,k}m_{j,k}.$$

   上述等式的左邊表示算子 M 作用於矩陣 T。而根據矩陣乘法的定義, 右邊正好等於 $\mathrm{Tr}\,(TM)$。由於 $F(T)$ 是實值的, 必然有 $m_{k,l}=\overline{m_{l,k}}$, 這表示 M 是自伴矩陣。

3. 我們說明: 一個自伴矩陣 M 是正定的, 若且唯若, 對於所有 $S\in\mathscr{D}_n$, 都有 $\mathrm{Tr}\,(SM)\geq 0$。這是因爲正定性的定義表示對於所有在 $\mathbb{C}^n$ 中的單位向量 $\vec{v}$, 都有 $\vec{v}\,M\left(\overline{\vec{v}}\right)^t\geq 0$。然而, 所有 “外積” $\langle\vec{v},\vec{v}\rangle$ 的集合是 $\mathscr{D}_n$ 的所有極點的集合 (根據上述性質)。

4. 最後, 我們考慮以下的總和: 由於對於每個狀態 S, $F(S)_u$ 是定義在 U 上的機率分佈 (我們回想測量的定義, 以及 U 是有限集合 $U = \{1, \ldots, k\}$), 因此滿足以下條件:

$$F(S)_i \geq 0,\ \forall\, i; \qquad \sum_i F(S)_i = 1.$$

根據步驟 3, 上述第一部分代表：對於所有 $i \in U$, 矩陣 M_i 是正定的。上述第二部分代表：對於所有 $S \in \mathscr{D}_n$, 有

$$\mathrm{Tr}\left(S\big(\sum_i M_i\big)\right) = 1,$$

但是, 此代表

$$\sum_i M_i = I_n,$$

其中 I_n 是單位矩陣。 Q.E.D.

結語: 在本文中, 我們論述了兩個機率和量子量測的關連。一個是推廣前文 [4] 中的 Max Born 法則; 一個是取材自 [3], 提出的量子量測的機率架構, 且證明此架構的一個基本結果。據作者所知, 此架構似乎迄無物理學者的迴響; 然則, 我們仍可以說, 此架構似乎很適合於有興趣於量子力學之數學理論的讀者進一步學習。尤其是備受重視的量子糾纏 (quantum entanglement) 理論, 本身就是和機率有密切關連, 而 [3] 中並沒有提到此主題, 只有在書末提到 Bell 不等式 (註: 原書俄文版是 1980 年出版) 。

参考資料

1. Sean Carroll, *Something Deeply Hidden: Quantum Worlds and the Emergence of Spacetime*, Random House, 2019.
2. Albert Einstein, Boris Podolsky and Nathan Rosen, Quantum-Mechanical Description of Physical Reality Be Considered Complete? *Physical Review*, 47(10), Princeton, New Jersey: Institute for Advanced Study, 777-780, 1935.
3. Alexander Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, Springer Nature, 2011.
4. 謝南瑞。機率學者看量子力學。數學傳播季刊, 48(4), 66-72, 2024。

—本文作者爲台灣大學數學系退休教授—